PCT WORLD ORGANIZATION OF INTELLECTUAL PROPERTY [seal]
International Office

APPLICATION PUBLISHED PURSUANT TO THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[6]: C12N 15/62, 15/14, 15/12, 15/60, 15/80, 15/81, 1/19, 1/15, A61K 38/38, 38/17, C07K 14/765// C12N 9/88, 1/21, 5/10

(11) International Publication No.: **WO 95/307,59**

(43) International publication date: November 16, 1995 (11/16/95)

A1

(21) International Application No.: PCT/FR95/00520

(22) International filing date: April 20, 1995 (4/20/95)

(30) Data relative to priority:
94/05616 May 6, 1994 (5/6/94) FR

(71) Applicant *(for all the designated states except US):* RHONE-POULENC RORER S.A. [FR/FR]; 20, avenue Raymond-Aron, F-92160 Antony (FR)

(72) Inventors; and

(75) Inventors/Applicants (*US only*): BECQUART, Jérôme [FR/FR]; 8, rue Maublanc, F-75015 Paris (FR). CONSEILLER, Emmanuel [FR/FR] 10, rue de Plélo, F-75015 Paris (FR). GUITTON, Jean-Dominique [FR/FR]; 74, rue Dunois, F-75013 Paris (FR). HARDY, Florence [FR/FR]; 6, square Villaret-de-Joyeuse, F-75017 Paris (FR). YEH, Patrice [FR/FR]; 11 bis, rue Lacépède, F-75005 Paris (FR).

(74) Agent: LE COUPANEC, Pascale; Rhône-Poulenc Rorer S.A., Patent Department, 20, avenue Raymond-Aron, F-92160 Antony (FR).

(81) Designated states: CA, JP, US, European Patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE)

Published
With international search report

(54) TITLE: BIOLOGICALLY ACTIVE POLYPEPTIDES INSERTED INTO AN ALBUMIN

(57) Abstract

Biologically active recombinant polypeptides essentially consisting of at least one active portion derived from a natural or artificial biologically active polypeptide and inserted into an albumin or albumin variant, the preparation thereof, and pharmaceutical compositions containing same, are disclosed.

Only by way of information

Codes used for designating PCT participating states listed on the cover sheets of international application documents published in accordance with the PCT

AT Austria
AU Australia
BB Barbados
BE Belgium
BF Burkina Faso
BG Bulgaria
BJ Benin
BR Brazil
BY Belarus
CA Canada
CF Central African Republic
CG Congo
CH Switzerland
CI Ivory Coast
CM Cameroon
CN China
CS Czechoslovakia
CZ Czech Republic
DE Germany
DK Denmark
ES Spain
FI Finland
FR France
GA Gabon
GB United Kingdom
GE Georgia
GN Guinea
GR Greece
HU Hungary
IE Ireland
IT Italy
JP Japan
KE Kenya
KG Kirghizistan
KP Korean Democratic People's Republic
KR Republic of Korea
KZ Kazakhstan
LI Liechtenstein
LK Sri Lanka
LU Luxembourg
LV Latvia
MC Monaco
MD Republic of Moldavia
MG Madagascar
ML Mali
MN Mongolia
MR Mauritania
MW Malawi
NE Nigeria
NL Netherlands
NO Norway
NZ New Zealand
PL Poland
PT Portugal
RO Romania
RU Russian Federation
SD Sudan
SE Sweden
SI Slovenia
SK Slovakia
SN Senegal
TD Chad
TG Togo
TJ Tajikistan
TT Trinidad and Tobago
UA Ukraine
US United States of America
UZ Uzbekistan
VN Vietnam

Biologically active polypeptides inserted into an albumin

The present invention concerns novel biologically active polypeptides, their preparation and pharmaceutical compounds containing them.

Although possessing one or more potential therapeutic activities, numerous polypeptides, unfortunately, cannot be pharmaceutically exploited. This can be for different reasons, such as, notably, their poor stability in vivo, their complex or fragile structure, the difficulty of producing them on an industrially acceptable scale, etc. In addition, certain polypeptides do not give the expected results in vivo due to problems of administration, packaging, pharmacokinetics, etc.

The object of the present invention is precisely to remedy these disadvantages.

It particularly involves the creation of biologically active artificial proteins permitting an optimal therapeutic exploitation of the biological properties of these polypeptides.

The Applicant has thus demonstrated that it is possible to insert by genetic engineering any active structure that is derived from a biologically active polypeptide into another protein structure made up of albumin, without altering said biological properties. In an unexpected manner, human serum albumin permits effectively presenting the active structure at its sites of interaction and assures a high plasma stability with the recombinant polypeptide of the invention.

More precisely, the present invention concerns a recombinant polypeptide having at least one active part derived from a natural or artificial, biologically active polypeptide, which is genetically inserted into an albumin or one of its variants or derivatives.

By albumin variant is meant, according to the present invention, any protein with a long plasma half-life that is obtained by modification, by means of genetic engineering techniques, of a gene coding for a given isomorph of human serum albumin, as well as any macromolecule with a long plasma half-life that is obtained by in vitro modification of the protein coded for by such genes. (By modification is meant any mutation, substitution, deletion, addition, or modification of the genetic and/or chemical nature). Since albumin is very polymorphic, numerous natural variants have been identified and cataloged [Weitkamp L.R., et al., Ann. Hum. Genet. 37 (1973) 219].

With regard to albumin derivatives, these most particularly involve molecules comprising all or part of the albumin, fused if necessary to at least one polypeptide sequence originating from a natural or artificial gene, itself endowed with biological activity or not so endowed.

In the following description, the different types of albumins explained above are commonly designated under the term albumin.

In the sense of the present invention, by active part is meant a part having an activity that can be either direct (treatment of illnesses, diagnosis, biological research, probes, etc.) or indirect (for example, usable in the prevention of illnesses, in the design of vaccines, in medical imaging techniques, etc.).

The active parts of biologically active polypeptides, inserted according to the invention, preferably are of therapeutic interest.

The polypeptides having a therapeutic activity can be of human or nonhuman origin.

By way of representation of polypeptides of nonhuman origin, one can cite peptides or their derivatives that have potentially useful properties in pathologies of the blood and interstitial compartments, such as hirudin, tigramin, antistatin, tick anticoagulant peptides (TAP), arietin, applagin, etc.

According to a preferred mode of the invention, the polypeptide having a therapeutic activity is a polypeptide of human origin or a molecular variant. For example, it can be all or part of an enzyme, an enzyme inhibitor, an antigen, an antibody, a hormone, a receptor, a factor involved in control of clotting, an interferon, a cytokine [the interleukins, but also their natural antagonist variants that prevent their binding to receptor(s), cytokines of the SIS (small induced secreted) type and, for example, macrophage inflammatory proteins (MIPs), etc.], a growth and/or differentiation factor [for example transforming growth factors (TGFs), blood-cell differentiation factors (erythropoietin, M-CSF, G-CSF, GM-CSF, etc.), insulin and insulin-like growth factors (IGFs), or even cell permeability factors (VPF/VEGF), etc.], factors involved in the genesis/resorption of bony tissues (OIF and osteospontin, for example), factors involved in cell motility or migration [and, for example, autocrine mobility factor (AMF), migration stimulation factor (MSF) or even a dispersion factor (scatter factor/hepatocyte growth factor)], a bactericidal or antifungal factor, a chemotactic factor [for

example, platelet factor 4 (PF4), or even monocyte chemoattractant peptides (MCP/MCAF) or neutrophilic chemoattractant factor, (NCAF), etc.], a cytostatic factor (for example, proteins that bond to galactosides), a plasma adhesion molecule (for example, von Willebrand factor, fibrinogen, etc.) or an interstitial (laminin, tenascin, vitronectin, etc.) or extracellular matrix molecule, or any other antagonist or agonist peptide sequence for molecular and/or intercellular interactions involved in pathologies of circulatory and interstitial compartments, for example the formation of arterial and venous thrombi, cancer metastases, tumoral angiogenesis, inflammatory shock, autoimmune disorders, osseous and osteoarticular pathologies, etc.

Of course, the active part of the polypeptides of the invention can be made up by the entire biologically active polypeptide or by a structure derived from the latter or can even correspond to a non-natural peptide sequence, which is isolated, for example, from any peptide bank. (For reasons of simplification, these different possibilities will be covered in the following under the common designation "active part of a biological active peptide"). In the sense of the present invention, derived structure means any polypeptide that is obtained by modification and that preserves biological activity. By modification is meant any mutation, substitution, deletion, addition or modification of the genetic and/or chemical nature. Such derivatives can be generated for different objectives, such as, in particular, increasing the affinity of the molecule for its binding sites, improving its production levels, increasing its resistance to proteases, increasing its therapeutic efficacy or even reducing its side effects, or conferring new

biological properties to it. By way of example, the chimeric polypeptides of the invention possess pharmacokinetic properties and a biological activity that can be used for the prevention or treatment of illnesses.

Particularly advantageous polypeptides of the invention are those in which the active part comprises:

(a) the entire peptide structure or

(b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or more residues) and having a therapeutic activity.

Among the structures of type (b), one can cite more particularly molecules in which certain sites of N- or O-glycosylation have been modified or removed, molecules in which one or more residues have been substituted, or molecules in which all the cysteine residues have been substituted. One can also cite molecules obtained from (a) by deletion of regions not involved or very little involved in the interaction with the bonding sites considered, or expressing an undesirable activity, and molecules having supplemental residues with regard to (a), such as, for example, an N-terminal methionine and/or a secretion signal and/or a junction peptide.

The subject of the present invention is particularly advantageous for active peptides that are too small to form a structural domain and/or that do not have a good in-vivo stability and/or a good bioavailability. The insertion proposed according to the invention permits combining them with one or more pre-existing

domains of the albumin and thus promoting the bioavailability and in-vivo stability of the latter.

In a general way, the size of the active part inserted into the albumin varies between three to twenty-five amino acid residues. However, sequences ranging from 1 residue to 100 residues can also be used.

The insertion of an active peptide into the peptide sequence of the albumin is conducted according to the invention in a manner so as to satisfy the following two conditions:

A sufficient accessibility must be preserved at said active part, inserted into the core of the albumin, in order to keep intact its biological activity. Moreover, the structure of the albumin cannot undergo too much destabilization, which would, of course, be detrimental to the recombinant polypeptide called the chimera.

Insertion sites are preferentially selected within the albumin by respecting the preceding imperatives.

According to the crystalline structure published by He and Carter (Nature 1992, 358, 209-215), albumin is formed by the repetition of 3 domains, each comprising two subdomains and it contains more than 67% alpha helices. Each of the domains can be superimposed over the others and is formed of 10 helices designated h1 to h10. Subdomain A has helices h1 to h6 and subdomain B has helices h7 to h10. Each subdomain is formed by a common motif: h1, h2, h3, h4 for domain A and h7, h8, h9 and h10 for domain B. Supplemental small helices

h5 and h6 are bound by a disulfide bridge to subdomain A. Figure 1 schematically shows the structure of human albumin.

The insertion sites are preferably localized in the regions of the albumin presumed to form exposed regions at the surface of the molecule, these regions preferentially being loops.

Insertion sites that are particularly suited to the invention are three regions of the first domain:

- region 5 which extends from residues 57 to 62 and which corresponds to a loop connecting the two helices h3 and h4;

- region 8 comprising residues 103 to 120 and corresponding to the zone between subdomains.

- region 13 comprised between residues 178 and 200 and which corresponds to a helix.

Helices h2 and h3 of domain III also define another insertion region from residue 419 to residue 430 that can be used according to the invention.

An active part of the biologically active peptide can be inserted in three different ways into the albumin peptide sequence:

- It can involve a strict insertion consisting of a simple addition of the peptide sequence of interest into the original sequence of the albumin which is preserved in its entirety.

- The insertion can correspond to a substitution of a portion of the peptide sequence of the albumin by the peptide sequence corresponding to the active part of the peptide of interest.

- Finally, it can be an insertion combining an addition of a part of the active peptide sequence and a substitution of a portion of the peptide sequence of the albumin by the rest of the active part of the active peptide.

Figure 2 schematically shows these different insertion methods.

Of course, the active part of a biologically active peptide can be repeated several times in the chimera at the same place and/or in different regions of the albumin. Moreover, it is also possible to insert different parts according to the invention, either from the same peptide or from different peptides.

Finally, an active part of the peptides according to the invention can be inserted either strictly inside the albumin or surrounded by junction sequences.

With regard to these junction sequences, they can notably be peptides rich in glycine residues and/or serine residues and/or threonine residues and/or any amino acid residue described as frequently encountered in flexibility zones of proteins.

The recombinant polypeptides of the invention have proven particularly advantageous.

They permit maintaining a given biological activity in the organism for a prolonged time. It has thus been proven possible to reduce the dosages administered and, in certain cases, to potentiate the therapeutic effect, for example, and simultaneously reduce the side effects that occur after administration of a larger amount. They advantageously permit generating and using structures that are derived from very small biologically active polypeptides and therefore are very specific for a desired effect. These recombinant

polypeptides also possess a particularly advantageous distribution in the organism, which changes their pharmacokinetic properties and favors the development of their biological activity and their use. They also have the advantage of being weakly immunogenic or nonimmunogenic for the organism in which they are used. Finally, the polypeptides of the invention can be expressed (and preferentially secreted) by recombinant organisms, at levels permitting their industrial exploitation.

Another subject of the invention concerns a preparation process for the chimeric molecules described above. More precisely, this process consists of causing a eukaryotic or prokaryotic host to express a nucleotide sequence coding for an active part of a desired polypeptide, and then harvesting the polypeptide produced.

Among eukaryotic hosts that can be used within the scope of the present invention, one can cite animal cells, yeasts or fungi. In particular, with regard to yeasts, one can cite yeasts of the genus Saccharomyces, Kluyveromyces, Pichia, Schwanniomyces or Hansenula. With regard to animal cells, one can cite COS, CHO, Cl27, etc. cells. Among fungi that can be used in the present invention can be cited more particularly Aspergillus ssp. or Trichoderma ssp. As prokaryotic hosts, it is preferred to use bacteria such as Escherichia coli, or those belonging to the genuses Corynebacterium, Bacillus or Streptomyces.

Nucleotide sequences that can be used within the scope of the present invention can also be prepared in different ways. Generally, they are obtained by assembling in reading phase the sequences coding for each of the functional

parts of the polypeptide. The latter can be isolated by the techniques of the person with average skill in the art, and ,for example, directly from cellular messenger RNA, (mRNA) or by recloning from a complementary DNA bank (cDNA), or they may even be totally synthetic nucleotide sequences. It is moreover understood that the nucleotide sequences can also be modified subsequently, for example by genetic engineering techniques, in order to obtain derivatives or variants of said sequences.

This nucleotide sequence coding for the active part of the polypeptide can be inserted directly or indirectly, depending on the region chosen for the insertion site, into the gene coding for albumin.

The selected region, in fact, may not have an adequate restriction site for conducting the insertion. In this hypothetical case, it may prove interesting, prior to insertion, to introduce one or more unique restriction sites. Restriction sites, at the level of the region chosen for the insertion site, are created preferably by directed mutagenesis according to classical techniques. However, one can also insert the nucleotide sequence corresponding to "the active part of the biologically active polypeptide" directly by directed mutagenesis without causing particular restriction sites to appear.

In the particular case of region 5 of the gene coding for albumin, the presence of the PVuII restriction site allows the direct insertion of the nucleotide sequence. It has proven to be particularly useful as a cloning site for an active part of a polypeptide that one wishes to insert in translational phase in the albumin sequence at the level of the 57^{th} residue.

In this mode of insertion, by profitably using a unique restriction site already existing in the original albumin sequence, the ligation of the sequence coding for the active peptide with the restriction fragment, which is linked to the entire gene coding for albumin, generates a nucleotide sequence comprising a hybrid gene coding for a chimeric protein of the strict HSA insertion type.

In the particular case of [residue] zone 419 to 430, the presence of 2 unique insertion sites (HincII and ArvII) permits inserting and/or substituting the peptides of biological interest in the albumin sequence.

With regard more particularly to regions 8 and 13, the insertion of said sequence is favored if one or more manipulatable restriction sites are created here. Of course, the restriction sites to be created at the sequence level are chosen by taking into account the nature of the restriction sites already existing. They must lead to a selective insertion.

In this second mode of embodiment, the use of two unique restriction sites permits constructing genes coding for chimeras containing the active peptide by means of insertion and/or by means of substitution. The peptide is inserted by means of replacing a fragment bordered by two unique restriction sites, in the complementary DNA of albumin, inserted by directed mutagenesis. These two unique restriction sites could be Mst I and Kpn I in region 8, and Sst I and Xho I in region 13. The creation of these restriction sites may or may not modify the polypeptide sequence of human albumin. The subsequent cloning of the active peptide in the coding phase in this albumin gene can involve the coding sequence of the peptide exclusively or can correspond to a combination of the

coding sequence of the peptide and the coding sequence of the fragment deleted from the albumin.

The present invention also seeks the protection of variants of nucleotide sequences coding for the corresponding albumin, i.e., integrating at least one non-natural unique restriction site, i.e., not present in the original sequence.

Advantageously, the creation of these restriction sites can then serve for the insertion of one or more active parts of biologically active polypeptide(s) into the mature protein.

More preferentially, in the process of the invention, the nucleotide sequence makes up part of an expression cassette comprising a transcription initiation region (promoter region) that permits, in the host cells, the expression of the nucleotide sequence placed under its control and coding for the polypeptides of the invention. This region can originate from promoter regions of genes strongly expressed in the host cell used, the expression being constitutive or controllable. In the case of yeasts, it can be a promoter of the gene for phosphoglycerate kinase (PKG), for glyceraldehyde 3-phosphate dehydrogenase (GPD), for lactase (LAC4), for enolases (ENO), for alcohol dehydrogenases, (ADH), etc. With regard to bacteria, it can be the promoter for the right or left genes of the lambda bacteriophage (P_L, P_R), or even promoters for genes of tryptophan operons (P_{trp}) or lactose operons (P_{lac}). Moreover, this control region can be modified, for example, by in vitro mutagenesis, by introduction of additional control elements or synthetic sequences, or by deletions or substitutions of the original control elements. The expression cassette can also

have a termination region for functional transcription in the envisioned host, positioned immediately downstream from the nucleotide sequence coding for a polypeptide of the invention.

In a preferred mode, the polypeptides of the invention result from expression in a eukaryotic or prokaryotic host of a nucleotide sequence and the secretion of the expression product of said sequence into the culture medium. It is in fact particularly advantageous to be able to obtain molecules directly in the culture medium by the recombinant method. In this case, the nucleotide sequence coding for a polypeptide of the invention is preceded by a "leader" sequence (or signal sequence) directing the polypeptide to the secretion pathways of the host used. This "leader" sequence can be the natural signal sequence of the biologically active polypeptide in the case where the latter is a naturally secreted protein, or that of the stabilizing structure, but it can also be any other functional "leader" sequence, or an artificial "leader" sequence. The choice of one or the other of these sequences is notably guided by the host used. Examples of functional signal sequences include those of sexual pheromone genes or yeast "killer" toxins.

In addition to the expression cassette, one or more markers that permit selecting the recombined host can be added, such as, for example, the URA3 gene of S. cerevisiae yeast, or genes conferring resistance to antibiotics such as geneticin (G418) or to any other toxic compound such as certain metal ions.

The assembly made up of the expression cassette and the selection marker can be introduced directly into the host cells considered, or first inserted

in a functional self-replicating vector. In the first case, sequences homologous to regions present in the genome of the host cells are preferentially added to this assembly; said sequences being then positioned on each side of the expression cassette and the selection gene so as to increase the integration frequency of the assembly into the genome of the host by targeting the integration of the sequences by homologous recombination. In the case where the expression cassette is inserted into a replication system, one that is preferred for yeasts of the genus Kluyveromyces is derived from plasmid pKD1 initially isolated from K. drosophilarum; a preferred replication system for yeasts of the Saccharomyces type is derived from plasmid 2μ of S. cerevisiae. Moreover, these expression plasmids may contain all or part of said replication systems, or may combine elements derived from plasmid pKD1 as well as plasmid 2μ.

Moreover, the expression plasmids can be shuttle vectors between a bacterial host such as Escherichia coli and the host cell chosen. In this case, a replication origin and a selection marker functioning in the bacterial host are required. It is also possible to position restriction sites surrounding bacterial sequences and that are unique on the expression vector: this permits eliminating these sequences by cutting and in vitro re-ligation of the cleaved vector before transformation of the host cells, which can result in an increase in the number of copies and an increased stability of the expression plasmids in said hosts. For example, such restriction sites can correspond to sequences such as 5'-GGCCNNNNNGGCC-3' (SfiI) (SEQ. ID No. 1) or 5'-GCGGCCGC-3' (NotI) (SEQ.

ID No. 2) to the extent that these sites are extremely rare and generally absent from an expression vector.

After construction of such vectors or expression cassettes, the latter are introduced into host cells according to the classical techniques described in the literature. For this purpose, any method permitting the introduction of a foreign DNA into the cell can be used. It can notably involve transformation, electroporation, conjugation, or any other technique known to the person of average skill in the art. By way of example for hosts of the yeast type, the different strains of Kluyveromyces used has been transformed by treating whole cells in the presence of lithium acetate and polyethylene glycol, according to the technique described by Ito et al. [J. Bacteriol. 153 (1983) 163]. The transformation technique described by Durrens et al. [Curr. Genet. 18 (1990) 7] using ethylene glycol and dimethyl sulfoxide has also been used. It is also possible to transform yeasts by electroporation, according to the method described by Karube et al. [FEBS Letters 182 (1985) 90]. An alternative protocol is also described in detail in the examples which follow.

After selection of transformed cells, the cells expressing said polypeptides are seeded and said polypeptides can be recovered, either during cell growth for "continuous" processes, or at the end of growth for "batch" ("batch") cultures. The polypeptides that make up the subject of the present invention are then purified from the culture supernatant in view of their molecular, pharmacokinetic and biological characterization.

A preferred expression system for the polypeptides of the invention consists of the use of yeasts of the genus Kluyveromyces as the host cell, transformed by certain vectors derived from the extrachromosomal replicon pKD1 initially isolated from K. marxianus var. drosophilarum. These yeasts, and in particular K. lactis and K. fragilis are generally capable of replicating said vectors in a stable manner and also have the advantage of being included in the list of G.R.A.S. ("Generally Recognized As Safe") organisms. Preferred yeasts are industrial strains of the genus Kluyveromyces capable of replicating in a stable manner said plasmids derived from plasmid pKD1 and in which a selection marker has been inserted as well as an expression cassette, permitting a high level of secretion of the polypeptides of the invention.

The present invention also concerns nucleotide sequences coding for the chimeric polypeptides described above, as well as the recombinant cells, eukaryotic or prokaryotic, containing such sequences.

The present invention also concerns the application by way of medication of the polypeptides according to the present invention. More particularly, the invention has for a subject any pharmaceutical composition containing one or more polypeptides or nucleotide sequences such as described above. These pharmaceutical compositions can be present in the form of diverse formulations. They can notably be nanoparticles on the surface of which the polypeptides according to the invention are present. This type of formulation is more particularly used to conduct directed targeting of the active ingredient. Of course, the nucleotide sequences can be used in gene therapy.

The present invention will be more completely described by means of the examples which follow, which must be considered as illustrative and non-limiting.

LIST OF FIGURES

The representations of plasmids given in the following Figures are not drawn to scale and only the restriction sites important for understanding the clonings conducted have been indicated.

Figure 1: Schematic representation of domain I of albumin with localization of the insertion sites according to the invention. * indicates the localization of insertion site h2-h3 in domain III. The numbers 5, 8 and 13 identify the corresponding insertion zones.

Figure 2: Schematic representation of different modes of insertion of an active peptide in the albumin structure.

Figure 3: pYG105 plasmid

Figure 4: Modification of region 5 of albumin after the insertion of the sequence coding for IEGR such as described in Example 8.1. The changes are shown in bold. The positions of the modified or introduced restriction sites are indicated by a horizontal line and the corresponding cutting position of the enzymes by a vertical line.

Figure 5: Cloning strategy for the insertion of IEGR in region 5 (Example 8.1).

Figure 6: Cloning strategy for the insertion of an active peptide into region 13 of the albumin.

EXAMPLES

GENERAL CLONING TECHNIQUES

The methods classically used in molecular biology such as preparative extractions of plasmid DNA, the centrifugation of plasmid DNA in cesium chloride gradient, agarose or acrylamide gel electrophoresis, purification of DNA fragments by electroelution, protein extraction with phenol or phenol-chloroform, precipitation of DNA in salt medium by ethanol or isopropanol, transformation in Escherichia coli, etc. are well known to the person of average skill in the art and are abundantly described in the literature [Maniatis T. et al., "Molecular Cloning, a Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982; Ausubel F.M. et al. (Eds.), "Current Protocols in Molecular Biology", John Wiley & Sons, New York, 1987].

The restriction enzymes were provided by New England Biolabs (Biolabs), Bethesda Research Laboratories (BRL) or Amersham and were used according to the recommendations of the manufacturers.

Plasmids of the pBR322, pUC type and phages of the M13 series are of commercial origin (Bethesda Research Laboratories).

For the ligation, DNA fragments are separated according to their size by agarose or acrylamide gel electrophoresis, extracted with phenol or by a phenol/chloroform mixture, precipitated with ethanol and then incubated in the presence of DNA ligase of T4 phage (Biolabs) according to the recommendations of the manufacturer.

The projecting 5' ends are filled in by the Klenow fragment of DNA Polymerase I from E. coli (Biolabs) according to the instructions of the manufacturer. The projecting 3' ends are eliminated in the presence of DNA Polymerase from phage T4 (Biolabs) used according to the recommendations of the manufacturer. The projecting 5' ends are eliminated by a treatment conducted with nuclease S1.

Directed mutagenesis in vitro by synthetic oligodeoxynucleotides is conducted according to the method developed by Taylor et al. [Nucleic Acids Res. 13 (1985) 8749-8764] by using the kit distributed by Amersham. This technique is notably implemented, within the scope of the present invention, to create unique restriction sites in view of a subsequent insertion.

The enzymatic amplification of DNA fragments by the technique called PCR [Polymerase-catalyzed Chain Reaction, Saiki R.K. et al., Science 230 (1985) 1350-1354; Mullis K.B. and Faloona F.A., Meth. Enzym. 155 (1987) 335-350] is conducted by using a "DNA thermal cycler" (Perkin Elmer Cetus) according to the instructions of the manufacturer.

The nucleotide sequences are verified by the method developed by Sanger et al. [Proc. Natl. Acad. Sci. USA, 74 (1977) 5463-5467] by using the kit distributed by Amersham.

The K. lactis transformations with DNA from the expression plasmids of the proteins of the present invention are conducted by any technique known to the person with average skill in the art, an example of which is given in the text.

Unless otherwise indicated, the bacterial strains used are E. coli MC1060 (lacPOZYA, X74, galU, galK, , strA^r), or E. coli TG1 (lac, proA, B, supE, thi, hsdD5/F'traD36, proA$^+$B$^+$, lal^q, lacZ, M15).

The yeast strains used belong to budding yeasts and more particularly to yeasts of the genus Kluyveromyces. The K. lactis MW98-8C strain (a, uraA, arg, lys, K$^+$, pKD1°) and K. lactis CBS 293.91 were particularly used; a sample of the MW98-8C strain was deposited on September 16, 1988, with the Centraalbureau voor Schimmelkulturen (CBS) in Baarn (the Netherlands) where it was registered under the number CBS 579.88.

A bacterial strain (E. coli) transformed with plasmid pET-8c52K was deposited on April 17, 1990, with the American Type Culture Collection under the number ATCC 68306.

The yeast strains transformed by the expression plasmids coding for the proteins of the present invention are cultured in Erlenmeyer flasks or 2-liter pilot fermenters (SETRIC, France) at 28°C in rich medium (YPD: 1% yeast extract, 2% Bactopeptone, 2% glucose; or YPL: 1% yeast extract, 2% Bactopeptone, 2% lactose) with constant stirring.

EXAMPLE 1:
PROTOCOL FOR STRICT PEPTIDE INSERTION USING A UNIQUE RESTRICTION SITE, PRESENT IN THE HSA GENE

By its uniqueness in the HSA gene and the associated vector, the PvuII site that is naturally localized in the coding sequence is particularly useful as the cloning site for a biologically active peptide that one wishes to insert in translation phase into HSA at the level of the 58th residue. In one particular mode of

embodiment, it is useful to use peptides of p residues whose coding sequence is [3xN]p. In this case, the oligonucleotides synthesized are of the 5'-NN [3xN]pN-3' type and its complementary strand. The ligation of this fragment with the HindIII-HindIII Δ PvuII restriction fragment, corresponding to the entire gene coding for HSA, generates a HindIII-HindIII restriction fragment comprising a hybrid gene coding for a chimeric protein of the strict insertion HSA type. In another mode of embodiment, the peptide can be repeated several times in the chimera.

EXAMPLE 2:
PROTOCOL FOR TOTAL OR PARTIAL INSERTION OF PEPTIDE BY USING TWO UNIQUE AND NATURAL RESTRICTION SITES

In one particular mode of embodiment, the use of two unique restriction sites permits constructing genes coding for chimeras having the active peptide by strict insertion or by substitution or by partial insertion.

For example, the existence of unique HincII and AvrII sites in the HSA coding sequence and in the vector permits generating a HindIII-HindIII Δ HincII-AvrII fragment. The elimination of the nucleotide fragment HincII-AvrII corresponds to the deletion of peptide fragment T(420) – N(429). The use of an appropriate complementary oligonucleotide permits cloning a peptide in coding phase in the HSA gene. This restriction fragment can be exclusively the complementary coding sequence of the peptide or can correspond to a combination of the coding sequence of the active peptide and the coding sequence of the fragment T(420) – N(429) of HSA. The active peptide can be present several times in the chimera.

EXAMPLE 3
EXPRESSION PLASMIDS

The chimeric proteins in the preceding examples can be expressed in yeasts from functional promoters, which are controllable or constitutive, such as, for example, those present in plasmids pYG105 (LAC4 promoter from Kluyveromyces lactis), pYG106 (PGK promoter from Saccharomyces cerevisiae), pYG536 (PHO5 promoter from S. cerevisiae) or hybrid promoters such as those described in Patent Application EP 361,991. Plasmids pYG105 and pYG106 are particularly useful here since they permit the expression of genes coded by the HindIII restriction fragments such as described in the preceding examples and cloned in the HindIII site and in productive orientation (described as the orientation which places the "prepro" region of the albumin proximal to the transcription promoter), starting from controllable functional promoters from K. lactis (pYG105) or constitutive promoters (pYG106). The pYG105 plasmid corresponds to the pKan707 plasmid described in Patent Application EP 361,991 in which the unique HindIII restriction site localized in the gene for resistance to geneticin (G418) was eliminated by directed mutagenesis while preserving an unchanged protein (oligodeoxynucleotide 5'-GAAATGCATAAGCTCTTGCCATTCTCACCG-3 (SEQ. ID No. 3). The SalI-SacI fragment coding for the URA3 gene of the mutated plasmid was then replaced by a SalI-SacI restriction fragment having an expression cassette made up of the LAC4 promoter from K. lactis (in the form of a SalI-HindIII fragment) and the terminator from the PGK gene of S. cerevisiae (in the form of a HindIII-SacI

fragment). The pYG105 plasmid is mitotitically very stable in Kluyveromyces yeasts. It is shown in Figure 3. The plasmids pYG105 and pYG106 differ only by the nature of the transcription promoter encoded by the SalI-HindIII fragment.

EXAMPLE 4:
TRANSFORMATION OF YEASTS

Yeasts belonging to the genus Kluyveromyces, and in particular the MW98-8C and CBS 293.91 strains of K. lactis, are transformed, for example, by the technique for treatment of whole cells by lithium acetate [Ito H. et al., J. Bacteriol. 153 (1983) 163-168], adapted as follows. The cells are grown at 28°C in 50 ml of YPD medium, with stirring and until an optical density of 600 nm (OD_{600}) is reached, comprised between 0.6 and 0.8; the cells are harvested by low-speed centrifugation, washed in a sterile solution of TE (10 mM Tris HCl, pH 7.4; mM EDTA), resuspended in 3-4 ml of lithium acetate (0.1 M in TE) in order to obtain a cell density of approximately 2×10^8 cells/ml, then incubated at 30°C for 1 hour with moderate stirring. Aliquots of 0.1 ml of the resulting suspension of competent cells are incubated at 30°C for 1 hour in the presence of DNA and at a final concentration of 35% polyethylene glycol (PEG_{4000}, Sigma). After a thermal shock of 5 minutes at 42°C, the cells are washed 2 times, resuspended in 0.2 ml of sterile water and incubated for 16 hours at 28°C in 2 ml of YPD medium to permit the phenotype expression of the gene for resistance to G418 expressed under the control of promoter P_{k1} (see EP 361,991); 200 µl of the cell suspension were then spread out on selective YPD boxes (G418, 200 µg/ml). The boxes are incubated at 28°C and the transformants appear after 2 to 3 days of cell growth.

EXAMPLE 5:
SECRETION OF CHIMERAS

After selection on rich medium supplemented with G418, the recombinant clones are tested for their capacity to secrete the mature form of the chimeric proteins. Several clones corresponding to the strain CBS 293.91 or MW98-8C transformed by the expression plasmids of the chimeras between HSA and the biologically active part are incubated in YPD or YPL medium at 28°C. The cell supernatants are recovered by centrifugation when the cells reach the stationary phase of growth, concentrated 10 times if needed by precipitation for 30 minutes at -20°C in a final concentration of 60% ethanol, then tested after SDS-PAGE () gel electrophosphoresis, or directly by gel staining with Coomassie blue, or after immunoblotting by using primary antibodies directed against the biologically active part or a polyclonal rabbit serum directed against HSA. During the immunological detection experiments, the nitrocellulose filter is first incubated in the presence of specific primary antibodies, washed several times, incubated in the presence of goat antibodies directed against the primary antibodies, and then incubated in the presence of an avidin-peroxidase complex by using the "ABC kit" distributed by Vectastain (Biosys S.A., Compiegne, France). The immunological reaction is then visualized by addition of 3,3-diaminobenzidine tetrahydrochloride (Prolabo) in the presence of hydrogen peroxide, according to the recommendations of the manufacturer.

EXAMPLE 6:
PURIFICATION OF CHIMERAS

The chimeras present in the culture supernatants corresponding to the transformed CBS 293.91 strain are characterized first by means of antibodies specific for the HSA part. It may be desirable to purify some of these chimeras. The culture is then centrifuged (10,000 g, 30 min), the supernatant is passed through a 0.22-mm filter (Millipore), then concentrated by ultrafiltration (Amicon) by using a membrane with a discrimination threshold of 30 kDa. The concentrate obtained is then dialyzed against a solution of Tris HCl (50 mM pH 8) and then purified on a column. For example, the concentrate corresponding to the culture supernatant of the transformed CBS 293.91 strain is purified by affinity chromatography on Trisacryl blue (IBF) and the samples are then dialyzed against water. A purification by molecular sieve may then be conducted. In this case, a Superose 12 column (Pharmacia) is equilibrated beforehand in a 20 mM NaH_2PO_4, 100 mM NaCl buffer, pH 7.0, and the samples from the affinity chromatography are loaded onto the column, harvested and characterized. A purification by ion exchange chromatography can also be used. In certain cases, the concentrate obtained after ultrafiltration is dialyzed against a solution of Tris HCl (50 mM, pH 8), then deposited by fractions of 20 ml onto a cation exchange column (5 ml) (S Fast Flow, Pharmacia), which has been equilibrated with the same buffer. The column is then washed several times with the Tris HCl solution (50 mM, pH 8) and the chimeric protein is then eluted from the column by an NaCl gradient (0 to 1 M). The fractions containing the chimeric protein are then combined and dialyzed against a solution of 50 mM Tris HCl (pH 8) and then

redeposited on an S Fast Flow column. After elution from the column, the fractions containing the protein are combined, dialyzed against water and lyophilized before characterization.

EXAMPLE 7:
INSERTION OF PEPTIDE 11 INTO THE ALBUMIN SEQUENCE

Peptide 11 has been described as the epitope of tryptophan synthase (Larvor et al., Mol. Immunol. (1991) 28, 523-531).

1. Strict insertion in a region naturally having unique restriction sites

Peptide 11 has the following peptide sequence in the direction from the N terminal to the C terminal: HGRVGIYFGMK (SEQ. ID No. 20). The insertion strategy in the HincII-AvrII zone consisted of ligating a synthesized nucleotide fragment coding for peptide 11 such that the reading frame for the coding sequence of HSA is respected and that the nature of the coding sequence of HSA is that of the original. We therefore first synthesized the following oligonucleotides: 5'-CCATGGTAGAGTAGGTATCTATTTCGGTATGAAAACTCCAACTCTTGTAGAG GTCTCGAGAAAT-3' (SEQ. ID No. 4) and 5'-CTAGATTTCTCGAGACCTCTACAAGATGTGGAGTTTTCATACCGAAATAGAT ACCTACTCTACCATGG-3' (SEQ. ID No. 5). These two oligonucleotides were hybridized and then ligated to the HindIII-HindIII Δ HincII-AvrII fragment, thus generating the entire HSA gene including the total insertion of peptide 11 between S(419) and T(420), immediately preceded by the "prepro" export region

of HSA. This fragment is cloned in productive orientation and in the HindIII site of plasmid pYG105.

2. Insertion by substitution and addition in a region naturally having unique restriction sites

In another mode of embodiment, the synthesized oligonucleotides were the following: 5'-CCATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3' (SEQ. ID No. 6) and 5'-CTAGTTTCATACCGAAATAGATACCTACTTCTTACCATGG-3' (SEQ. ID No. 7). These two oligonucleotides are hybridized and then ligated with the HindIII-HindIII Δ HincII-AvrII fragment, thus generating the gene of a chimera amputated from the residues T(420) to N(429) and substituted by the sequence of peptide 11. This fragment is cloned in productive orientation and in the HindIII site of plasmid pYG105.

3. Strict insertion in a region not having unique restriction sites

In another mode of embodiment, two unique restriction sites, Sst I and Xho I were created by directed mutagenesis. The same type of directional cloning was conducted. In the case of the total insertion of peptide 11 between residues A (191) and S (192), the two following oligonucleotides were synthesized: 5'-ACGGGATGAAGGGAAGGCCCATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3' (SEQ. 1D No. 8) and 5'-TCGATTTCATTACCGAAATAGATACCTACTCTACCATGGGCCTTCCCTTCATCCCGTAGCT-3' (SEQ. ID No. 9). We then proceeded, according to the protocol

already described in preceding paragraphs 1 or 2, to the creation of the corresponding expression plasmid.

4. Insertion by substitution in a region not having unique restriction sites

This insertion requires the creation beforehand of two restriction sites Sst 1 and Xho 1. For partial insertion per se, the two following nucleotides were synthesized: 5"-ACATGGTAGAGTAGGTATCTATTTCGGTATGAAA-3" (SEQ. ID No. 10 and 5'-TCGATTTCATACCGAAATAGATACCTACTCTACCATGTAGCT-3' (SEQ. ID No. 11). In this latter case, residues R (186) to A (191) were deleted and substituted by peptide 11. The two expression plasmids of these chimeras are 1671 and 1667, respectively.

Each of the plasmids, obtained at points 1, 2, 3 and 4, is used to transform a yeast strain according to the protocol described in Example 4; the corresponding proteins are secreted and purified according to Examples 5 and 6.

EXAMPLE 8
INSERTION OF THE IEGR PEPTIDE SEQUENCE, SUBSTRATE FOR FACTOR Xa

The peptide containing the sequence IEGR (SEQ. ID No. 21) is then inserted, the target of protease factor Xa, which converts prothrombin into thrombin in the reaction cascade occurring in blood clotting.

1. Strict insertion in region 5 having a unique PvuII site

In order to make this insertion at the PvuII site of the albumin gene, first a replicative vector devoid of the PvuII site is created, into which is then inserted

the gene coding for the albumin prepro sequence. The following two complementary oligonucletides, coding for the IEGR sequence, were also synthesized: 5'GATCCATAGAAGGTCGACTAG-3' (SEQ. ID No. 12) and 3'-CTAGGTATCTTCCAGCTGATC-5' (SEQ. ID No. 13). These two oligonucleotides are then hybridized and inserted at the level of the PvuII site of the albumin gene. The modifications that their insertion introduces in the nucleotide and peptide sequences of albumin are reported in Figure 4. The cloning strategy is diagrammed in Figure 5.

This construction illustrates the case where junction sequences are introduced on either side of the peptide of interest.

2. Insertion in region 13

Figure 6 describes a cloning strategy for a peptide in region 13.

In the following, restriction sites Sst I and Xho I were created by directed mutagenesis. In the case of region 13, we conducted, respectively, a total substitution and an addition substitution of the IEGR sequence alone or framed by junction sequences.

With regard to total substitution, the oligonucleotides used are the following:

5'-CAGAATCGAAGGTAGAGCC-3' (SEQ. ID No. 14 and 5'-TCGAGGCTCTACCTTCGATCGAGGGTAGCT-3 (SEQ. ID No. 15).

At the protein level, the sequence (187) DEGK (SEQ. ID No. 22) is substituted by IEGR.

In the second case, the oligonucleotides used are the following: 5'-ACCCTCGATCGAAGGTAGATCTCCA-3' (SEQ. ID No. 16), 5'-TCGATGGAGATCTACCTTCGATCGAGGGTAGCT-3' (SEQ. ID No. 17).

At the protein level, the albumin is amputated from residue R(186) to residue A(191) and replaced by the PSIEGRSP sequence (SEQ. ID No. 23), therefore leading to an addition of two residues.

The corresponding proteins are secreted and purified according to the protocols described in the preceding examples. Table I below summarizes the characteristics of these chimeras.

Table I

Albumin incorporating	Expression amount ug/l	Purification yield	Final purity
IEGR	150	15	98
PSIEGRPS	110	36	96

3. Biological activity of the chimera obtained according to Example 8.1

The chimera is incubated in the presence of bovine factor Xa in an enzyme/substrate ratio of 1/10 and in a 50 mM Tris, 100 mM NaCl, 1 mM $CaCl_2$ buffer, pH 8.0, for 3 hours at 37°C. After this treatment, an SDS-PAGE analysis is conducted which shows the cleavage of the chimera characterized by the generation of an albumin fragment of a mass of the order of 60 kD.

LIST OF SEQUENCES

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: RHONE-POULENC RORER S.A.

(B) STREET ADDRESS: 20, avenue Raymond ARON

(C) CITY: ANTONY

(E) COUNTRY: FRANCE

(F) POSTAL CODE: 92165

(ii) TITLE OF THE INVENTION: New biologically active polypeptides, their preparation and pharmaceutical composition containing them.

(iii) NUMBER OF SEQUENCES: 24

(iv) FORM READABLE BY COMPUTER

(A) TYPE OF SUPPORT [MEDIUM]: tape

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ. ID NO. 1:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 13 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 1:

GGCCNNNNNG GCC 13

(3) INFORMATION FOR SEQ. ID NO: 2:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 8 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 2:

GCGGCCGC 8

(4) INFORMATION FOR SEQ. ID NO: 3:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 3:

GAAATGCATA AGCTCTTGCC ATTCTCACCG 30

(5) INFORMATION FOR SEQ. ID NO: 4:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 64 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 4:

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAAACTCCA ACTCTTGTAG

AGGTCTCGAG AAAT 64

(6) INFORMATION FOR SEQ. ID NO: 5:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 68 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 5

CTAGATTTCT CGAGACCTCT ACAAGATGTG GAGTTTTCAT ACCGAAATAG

ATACCTACTC TACCATGG 68

(3) INFORMATION FOR SEQ. ID NO: 6:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 6:

CCATGGTAGA GTAGGTATCT ATTTCGGTAT GAAA 34

(8) INFORMATION FOR SEQ. ID NO: 7:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 40 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 7:

CTAGTTTCAT ACCGAAATAG ATACCTACTT CTTACCATGG 40

(9) INFORMATION FOR SEQ. ID NO: 8:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 52 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 8:

ACGGGATGAA GGGAAGGCCC ATGGTAGAGT AGGTATCTAT

TTCGGTATGA AA 52

(10) INFORMATION FOR SEQ. ID NO: 9:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 61 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 9:

TCGATTTCAT TACCGAAATA GATACCTACT CTACCATGGG CCTTCCCTTC

ATCCCGTAGC T 61

(11) INFORMATION FOR SEQ. ID NO: 10

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 10

ACATGGTAGA GTAGGTATCT ATTTCGGTAT GAAA 34

(12) INFORMATION FOR SEQ. ID NO: 11:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 42 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 11:

TCGATTTCAT ACCGAAATAG ATACCTACTC TACCATGTAG CT 42

(13) INFORMATION FOR SEQ. ID NO: 12:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 12:

GATCCATAGA AGGTCGACTA G 21

(14) INFORMATION FOR SEQ. ID NO: 13:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 13:

CTAGGTATCT TCCAGCTGAT C 21

(15) INFORMATION FOR SEQ. ID NO: 14:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 14:

CAGAATCGAA GGTAGAGCC 19

(16) INFORMATION FOR SEQ. ID NO: 15:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 15

TCGAGGCTCT ACCTTCGATC GAGGGTAGCT 30

(17) INFORMATION FOR SEQ. ID NO: 16:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 16:

ACCCTCGATC GAAGGTAGAT CTCCA 24

(18) INFORMATION FOR SEQ. ID NO: 17:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 17:

TCGATGGAGA TCTACCTTCG ATCGAGGGTA GCT 33

(19) INFORMATION FOR SEQ. ID NO: 18:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 18:

GTCCCGGATG GAGCGCGTAC TTAGAGAGAA T 31

(20) INFORMATION FOR SEQ. ID NO: 19:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) NUMBER OF STRANDS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTISENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 19:

TCGACAGGGC CTACCTCGCG CATGAATCTC TCTTAAGCT 39

(21) INFORMATION FOR SEQ. ID NO: 20:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 20:

His Gly Arg Val Gly Ile Tyr Phe Gly Met Lys
5 10

(22) INFORMATION FOR SEQ. ID NO: 21:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 21:

Ile Glu Gly Arg

(23) INFORMATION FOR SEQ. ID NO: 20:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 22:

Asp Glu Gly Lys

(24) INFORMATION FOR SEQ. ID NO: 23:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(D) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 23:

Pro Ser Ile Glu Gly Arg Ser Pro
5

(25) INFORMATION FOR SEQ. ID NO: 24:

(i) CHARACTERISTICS OF THE SEQUENCE:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(D) CONFIGURATION: linear

(xi) DESCRIPTION OF THE SEQUENCE: SEQ. ID NO: 24:

Arg Met Glu Arg Val Leu Arg Glu Asn
5

CLAIMS

1. Recombinant polypeptide having at least one active part derived from a natural or synthetic polypeptide, biologically active, genetically inserted into an albumin or one of its variants or derivatives.

2. Polypeptide according to claim 1, further characterized in that the biologically active polypeptide has a therapeutic activity and is of human origin.

3. Polypeptide according to claim 2, further characterized in that the polypeptide having a therapeutic activity is chosen completely or partially from enzymes, enzyme inhibitors, antigens, antibodies, hormones, receptors, clotting factors, interferons, cytokines, growth and/or differentiation factors, factors involved in the genesis/resorption of osseous tissues, chemotactic factors, factors for cellular motility or migration, cytostatic factors, bactericidal or antifungal factors, or adhesive plasma, interstitial, or extracellular matrix molecules.

4. Polypeptide according to one of claims 1 to 3, further characterized in that the polypeptide having a therapeutic activity is chosen from among any antagonist or agonist peptide sequence for molecular interactions and/or cellular interactions involved in pathologies of the circulatory and interstitial compartments.

5. Polypeptide according to any one of claims 1 to 4, further characterized in that the active part has a structure chosen from among:

(a) the whole peptide structure or

(b) a fragment of (a) or a structure derived from (a) by structural modification (mutation, substitution, addition and/or deletion of one or more residues) while conserving a therapeutic activity.

6. Polypeptide according to one of claims 1 to 5, further characterized in that the active part is inserted strictly inside the albumin or surrounded by junction sequences.

7. Polypeptide according to one of claims 1 to 6, further characterized in that the active part is inserted preferably at the level of the regions of the albumin presumed to form exposed regions at the surface of the molecule.

8. Polypeptide according to one of claims 1 to 7, further characterized in that the active part is inserted at the level of region 5 extending from residues 57 to 62 of the albumin.

9. Polypeptide according to one of claims 1 to 7, further characterized in that the active part is inserted at the level of region 8 extending from residues 103 to 120 of the albumin.

10. Polypeptide according to one of claims 1 to 7, further characterized in that the active part is inserted at the level of region 13 extending from residues 178 to 200 of the albumin.

11. Polypeptide according to one of claims 1 to 7, further characterized in that the active part is inserted at the level of the region from residue 415 to residue 425 defined by helices h2 and h3 of domain III in the albumin.

12. Polypeptide according to one of claims 1 to 12, further characterized in that the active part is inserted therein in a single or multiple form.

13. Polypeptide according to claim 12, further characterized in that the active part is repeated several times in the same place and/or in different regions of the albumin.

14. Polypeptide according to one of claims 1 to 13, further characterized in that the active parts inserted are of different nature.

15. Variant of a nucleotide sequence coding for albumin or one of its variants or derivatives integrating at least one unique non-natural restriction site.

16. Nucleotide sequence coding for a polypeptide according to any one of claims 1 to 14.

17. Nucleotide sequence according to claim 16, further characterized in that it comprises a "leader" sequence permitting the secretion of the polypeptide expressed.

18. Expression cassette comprising a nucleotide sequence according to one of claims 16 or 17 under the control of a transcription initiation region and possibly a transcription termination region.

19. Self-replicating plasmid having an expression cassette according to claim 18.

20. Recombinant eukaryotic or prokaryotic cell in which a nucleotide sequence according to either one of claims 16 or 17 or an expression cassette according to claim 18 or a plasmid according to claim 19 has been inserted.

21. Recombinant cell according to claim 20, further characterized in that it is a yeast, an animal cell, a fungus or a bacterium.

22. Recombinant cell according to claim 21, further characterized in that it is a yeast.

23. Recombinant cell according to claim 22, further characterized in that it is a yeast of the genus Saccharomyces or Kluyveromyces.

24. Process for preparation of a polypeptide such as defined in one of claims 1 to 14, characterized in that a recombinant cell according to one of claims 20 to 23 is cultured under expression conditions, and the polypeptide produced is recovered.

25. Pharmaceutical composition containing one or more polypeptides according to any one of claims 1 to 14.

26. Pharmaceutical composition containing a nucleotide sequence according to either one of claims 16 to 17 usable in gene therapy.

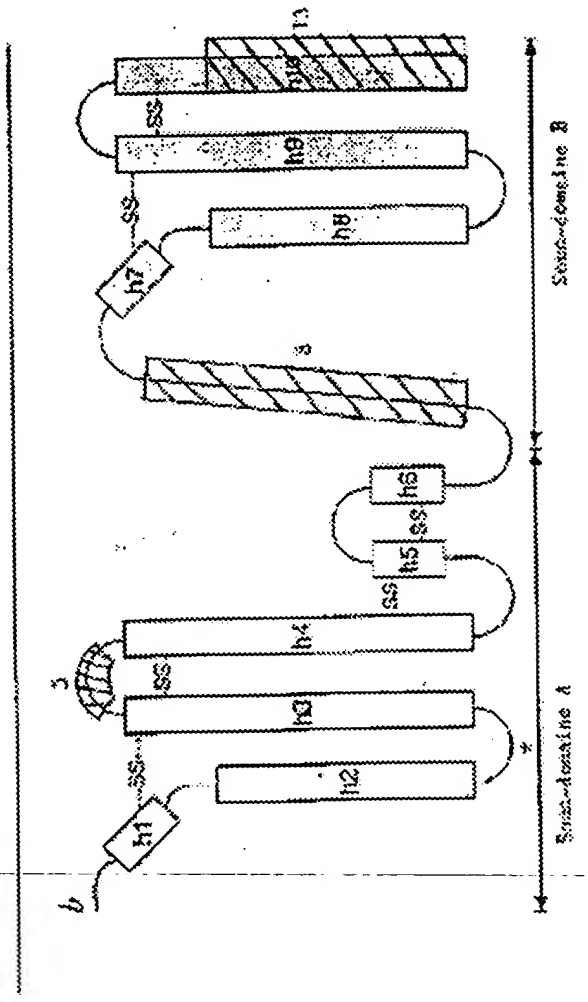


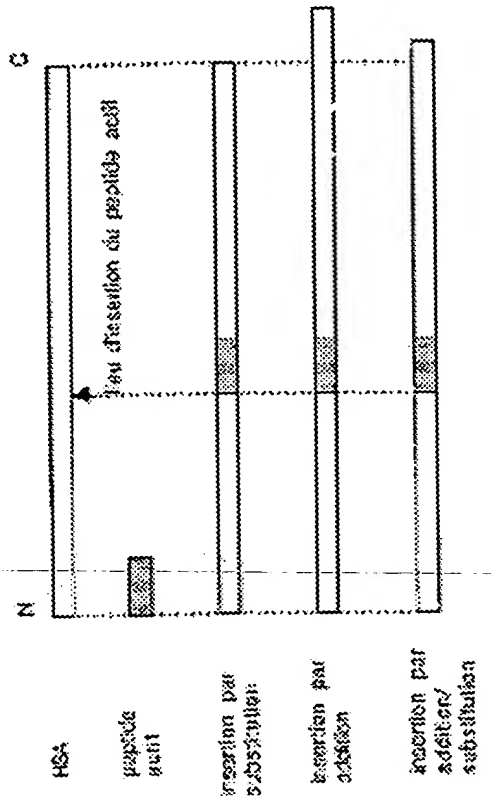
Figure 1

Sous-domaine = Subdomain

Fig. 1

Figure 2
left:
active peptide
insertion by substitution
insertion by addition
insertion by addition/substitution

right: insertion site for the active peptide



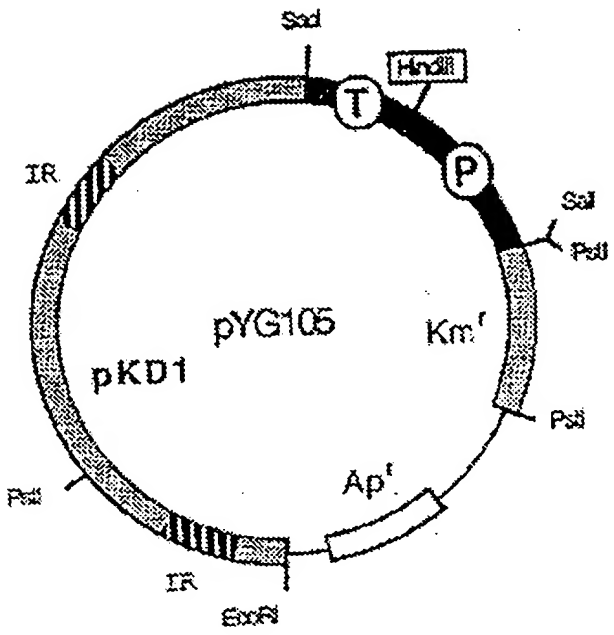

SacI
HindIII
T
P
SalI
PstI
IR
pYG105
pKD1
Km^r
PstI
Ap^r
PstI
IR
EcoRI

Figure 3

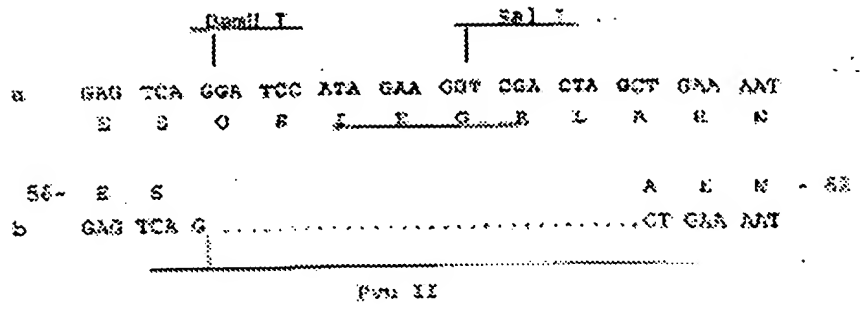

PvuII I
SalI I
a GAG TCA GGA TCC ATA GAA GGT CGA CTA GCT GAA AAT
E S G S I E G R L A E N
56- E S A E N - 53
b GAG TCA G CT GAA AAT
PvuII II

Figure 4

[handwritten]: Transcription of the prepro HSA
SAH = HAS
Ligature = ligation

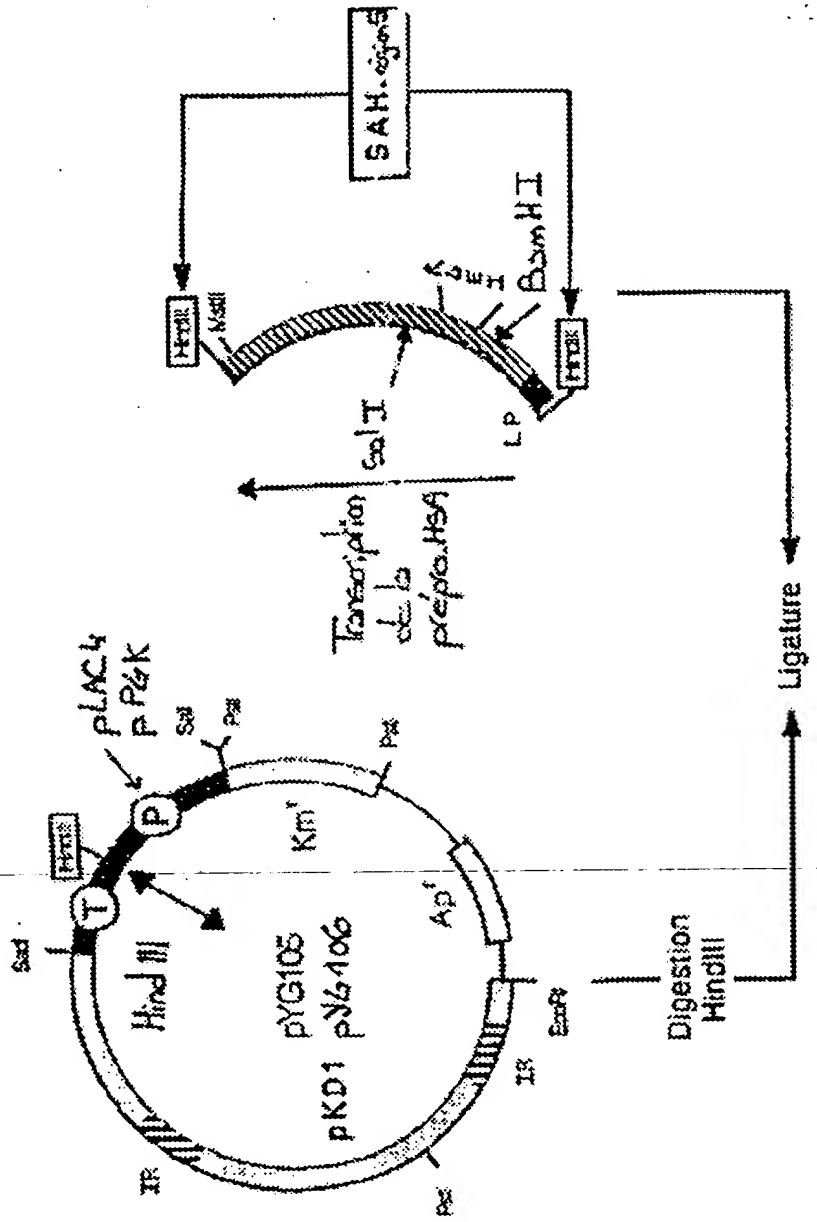
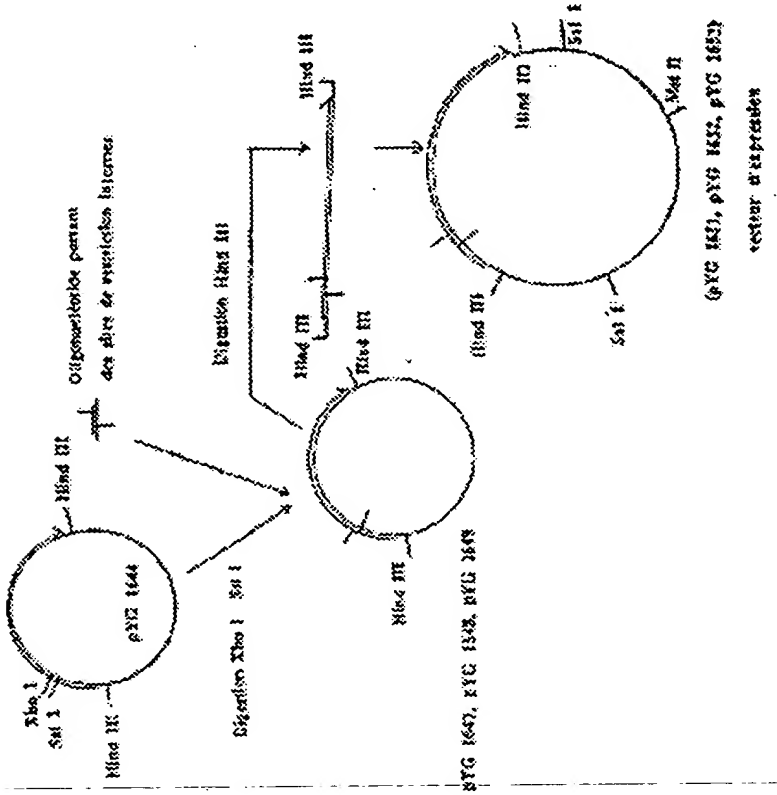


Figure 5

Figure 6

top right: Oligonucleotide bearing internal restriction sites
bottom right: Expression vector



[International Search Report is given in English]